



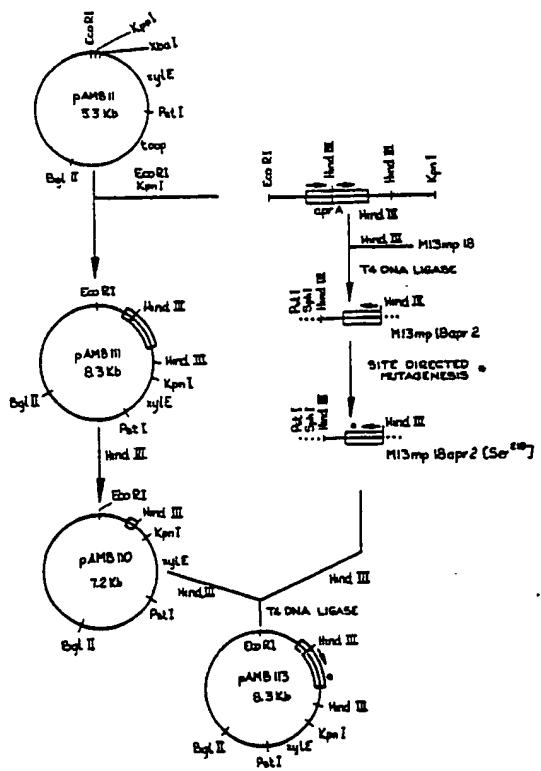
## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>4</sup> : C12N 9/56, C07H 15/12 C12P 21/02, 15/00, C12N 1/20		A1	(11) International Publication Number: WO 87/ 04461 (43) International Publication Date: 30 July 1987 (30.07.87)		
(21) International Application Number: PCT/US87/00027		(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), NO, SE (European patent).			
(22) International Filing Date: 7 January 1987 (07.01.87)					
(31) Priority Application Number: 819,241					
(32) Priority Date: 15 January 1986 (15.01.86)					
(33) Priority Country: US					
(71) Applicant: AMGEN [US/US]; 1900 Oak Terrace Lane, Thousand Oaks, CA 91320 (US).					
(72) Inventors: STABINSKY, Yitzhak ; 3415 Heidelberg Drive, Boulder, CO 80300 (US). ZUKOWSKI, Mark, M. ; 2928 Sunflower Street, Thousand Oaks, CA 91360 (US).					
(74) Agent: BORUN, Michael, F.; Marshall, O'Toole, Gerstein, Murray & Bicknell, Two First National Plaza, Suite 2100, Chicago, IL 60603 (US).					

(54) Title: THERMALLY STABLE AND pH STABLE SUBTILISIN ANALOGS AND METHOD FOR PRODUCTION THEREOF

(57) Abstract

A mutated subtilisin suitable for admixture to washing compositions and exhibiting substantially improved stability over naturally occurring *Bacillus* serine proteases is prepared by expressing a modified gene encoding subtilisin in *Bacillus subtilis*. A preferred subtilisin analog product differs from wild-type *Bacillus* alkaline proteases by having any amino acid, and preferably serine, at position (218) in place of asparagine. The product is preferably produced in a strain of *B. subtilis* which is mutated to block synthesis of endogenous proteases. The method of replacing an Asn or Gly in an Asn-Gly sequence in order to improve pH and thermal stability may be applied to other sites in subtilisin and to other proteins as well.



***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	ML	Mali
AU	Australia	GA	Gabon	MR	Mauritania
BB	Barbados	GB	United Kingdom	MW	Malawi
BE	Belgium	HU	Hungary	NL	Netherlands
BG	Bulgaria	IT	Italy	NO	Norway
BJ	Benin	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	LI	Liechtenstein	SN	Senegal
CH	Switzerland	LK	Sri Lanka	SU	Soviet Union
CM	Cameroon	LU	Luxembourg	TD	Chad
DE	Germany, Federal Republic of	MC	Monaco	TG	Togo
DK	Denmark	MG	Madagascar	US	United States of America
FI	Finland				

- 1 -

THERMALLY STABLE AND pH STABLE SUBTILISIN  
ANALOGS AND METHOD FOR PRODUCTION THEREOF

5

Background of the Invention

10 The present invention relates in general to thermally stable and pH stable analogs of the enzyme subtilisin and to a method for generating such analogs. In particular, the present invention relates to analogs of Bacillus subtilisin having a substitution 15 for Asn<sup>218</sup> and to a method for generating such analogs.

The term subtilisin designates a group of extracellular alkaline serine proteases produced by various species of Bacilli. These enzymes are also called Bacillus serine proteases, Bacillus subtilisins 20 or bacterial alkaline proteases.

The Bacillus subtilisin molecules are composed of a single polypeptide chain of either 274 residues (for subtilisin type Carlsberg produced by Bacillus licheniformis and for the subtilisin produced by 25 Bacillus subtilis strain DY) or 275 residues (for subtilisin type BPN', produced by Bacillus amyloliquefaciens, and the aprA gene product of Bacillus subtilis). When comparing amino acid sequences of subtilisin from different strains of Bacillus the sequence 30 of subtilisin BPN' is used as a standard. For example, based on an alignment of sequences that gives the highest degree of homology between subtilisin Carlsberg and subtilisin BPN', the serine at the active site of the former is referred to as serine 221, even though it 35 is located at position 220 of the amino acid sequence. On the same basis, position 220 of the amino acid

- 2 -

sequence of subtilisin Carlsberg may said to "correspond" to position 221 of subtilisin BPN'. See e.g., Nedkov et al., Hoppe-Seyler's Z. Physiol. Chem., 364, 1537-1540 (1983).

5        The X-ray structure of subtilisin BPN' [Wright, et al., Nature, 221, 235 (1969)] revealed that the geometry of the catalytic site of subtilisin, involving Asp<sup>32</sup>, His<sup>64</sup> and Ser<sup>221</sup>, is almost identical to that of the active site of mammalian serine proteases 10 (e.g., chymotrypsin) involving the residues Asp<sup>102</sup>, His<sup>57</sup>, and Ser<sup>195</sup>. However, the overall dissimilarities between Bacillus serine proteases and mammalian serine proteases indicate that these are two unrelated families of proteolytic enzymes.

15       In the family of Bacillus subtilisins complete amino acid sequences are available for four subtilisins: Carlsberg, [Smith, et al., J. Biol. Chem., 243, 2184-2191 (1968)]; BPN' [Markland, et al., J. Biol. Chem., 242, 5198-5211 (1967)]; the aprA gene product 20 [Stahl, et al., J. Bacteriol., 158, 411-418 (1984)]; and DY [Nedkov, et al., supra]. Subtilisin Carlsberg and subtilisin BPN' (sometimes referred to as subtilisin Novo) differ by 84 amino acids and one additional residue in BPN' (subtilisin Carlsberg lacks an amino 25 acid residue corresponding to residue 56 of subtilisin BPN'). Smith, et al., supra. Subtilisin DY is 274 amino acids in length and differs from subtilisin Carlsberg in 32 amino acid positions and from subtilisin BPN' by 82 amino acid replacements and one deletion 30 (subtilisin DY lacks an amino acid residue corresponding to residue 56 of subtilisin BPN'). Nedkov, et al., supra. The amino acid sequence of the aprA gene product is 85% homologous to the amino acid sequence of subtilisin BPN'. Stahl, et al., supra. Thus, it seems 35 that there is an extensive homology between amino acid sequences of serine proteases from different strains of

- 3 -

Bacillus. This homology is complete in certain regions of the molecule and especially in those that play a role in the catalytic mechanism and in substrate binding.

Examples of such sequence invariances are the primary and secondary substrate binding sites, Ser<sup>125</sup>-Leu<sup>126</sup>-Gly<sup>127</sup>-Gly<sup>128</sup> and Tyr<sup>104</sup> respectively and the sequence around the reactive serine (221) Asn<sup>218</sup>-Gly<sup>219</sup>-Thr<sup>220</sup>-Ser<sup>221</sup>-Met<sup>222</sup>-Ala<sup>223</sup>.

Subtilisin molecules exhibit some unique stability properties. They are not completely stable at any pH value although they are relatively resistant to denaturation by urea and guanidine solutions and enzymatic activity is retained for some time even in a solution of 8 M urea. In solutions at a pH below 4, subtilisin rapidly and irreversibly loses its proteolytic activity. Gounaris, et al., Compt. Rend. Trav. Lab. Carlsberg, 35, 37 (1965) demonstrated that the acid deactivation of subtilisin is not due to a general charge effect and speculated that it is due to other changes in the molecule, such as protonation of histidine residues in the interior, hydrophobic parts of the molecule. In solution at pH above 5, Bacillus serine proteases gradually undergo irreversible inactivation at a rate that increases with temperature and pH. The mechanisms of this inactivation are not fully known but there is evidence indicating that autodigestion is responsible at least in part for enzyme instability at this pH range.

The use of proteases in industrial processes which require hydrolysis of proteins has been limited due to enzyme instability under operational conditions. Thus, for example, the incorporation of trypsin into laundry detergents (e.g., Bio-38, Schnyder, Switzerland) to facilitate removal of proteinaceous stains had a very limited success which was undoubtedly a result of enzyme instability under the washing

- 4 -

conditions. It was only about 1960, after the introduction of the use of bacterial alkaline proteases which are more compatible with detergents that proteases came to be widely used in the detergent industry.

5 For practical reasons many industrial processes are conducted at temperatures that are above the stability range of most enzymes. Therefore, it is reasonable to assume that highly thermostable proteases not only will be advantageous to certain industries such 10 as detergent and hide dehairing, that already require stable proteases, but may be useful in industries that use chemical means to hydrolyze proteins e.g. hydrolysis of vegetable and animal proteins for the production of soup concentrates.

15 It should be pointed out, however, that although thermal inactivation may be the most important mode of enzyme inactivation, factors other than heat such as extremes of pH, oxygen and denaturing agents may have a determinantal effect on limiting the use of 20 proteases in industrial processes. It is therefore, desirable to obtain proteases that are characterized by improved stability under the operational conditions used in various industries. Such a goal may be accomplished either through searching for new more stable wild-type 25 enzymes or through stabilization of already known existing proteases.

Even though the Bacillus-derived alkaline proteases are more compatible with detergent formulations than were the pancreatic proteases, they are still 30 not ideal in all respects.

Over the past several years there have been major changes in detergent formulations, particularly in the replacement of phosphates with alternate builders and in the development of liquid laundry detergents to 35 meet environmental and consumer demands. These changes create a need for changes in traditional detergent

- 5 -

enzymes. More particularly, it has become desirable to employ proteolytic enzymes which possess greater storage stability in liquid laundry formulations as well as stability and activity at broader ranges of pH and 5 temperature.

In one approach to producing modified subtilisins for use in detergent formulations, as disclosed in European Patent Application No. 130,756, mutations in 10 the subtilisin of Bacillus amyloliquefaciens (B. amyloliquefaciens) at Tyr<sup>-1</sup>, Asp<sup>32</sup>, Asn<sup>155</sup>, Tyr<sup>104</sup>, Met<sup>222</sup>, Gly<sup>166</sup>, His<sup>64</sup>, Gly<sup>169</sup>, Phe<sup>189</sup>, Ser<sup>33</sup>, Ser<sup>221</sup>, Tyr<sup>217</sup>, Glu<sup>156</sup>, and/or Ala<sup>152</sup> are identified as providing changed stability, altered conformation or as 15 having changes in the "processing" of the enzyme. In this context, mutation of Met<sup>222</sup> to Ala, Cys (which mutant also exhibits a sharper pH optimum than wild type) or Ser assertedly results in improved oxidation stability. Substitution for Gly<sup>166</sup> with Ala, Asp, Glu, 20 Phe, Hys, Lys, Asn, Arg or Val appears to alter the kinetic parameters of the enzyme. However, none of the mutations are disclosed to provide analogs having greater stability at high temperatures or stability over a broader pH range than the wild type enzyme.

25 In another approach, it appears that the pH dependence of subtilisin may be altered, as disclosed in Thomas, et al, Nature, 318, 375-376 (1985), by changing an Asp to Ser in Asp<sup>99</sup>-Gly<sup>100</sup> of subtilisin BPN'. This change represents an alteration of a surface charge 14-30 15 Angstroms from the active site. However, the approach of Thomas, et al. does not provide an indication of improvement where no change in surface charge is made, as is the case where one uncharged residue is substituted for another.

- 6 -

Summary of the Invention

The present invention provides analogs of Bacillus serine protease products characterized by improved pH and heat stabilities, rendering them especially useful in detergent formulations as well as other processes requiring protease usage. The present invention also provides an industrially feasible recombinant process for preparing the products free of other proteases. Stable analogs according to the present invention are generally characterized by deletion and/or modification and/or replacement of either residue of Asn-Gly sequences present in the protease.

It should be noted that, as employed herein, the term "subtilisin" is used to refer to the mature, secreted form of the enzyme which lacks leader sequences cleaved from the mature enzyme prior to or at secretion.

Presently preferred analogs of a Bacillus subtilisin according to the present invention have an amino acid sequence wherein positions comprising an Asn-Gly sequence in the Bacillus subtilisin do not comprise an Asn-Gly sequence in the analog, and in particular wherein there are fewer Asn-Gly sequences than in the Bacillus subtilisin. Most preferably, a position corresponding to position 218 in the amino acid sequence as set forth in Table 1, does not comprise an asparaginyl residue, but rather comprises a residue of a different amino acid, preferably an amino acid selected from among serine, valine, threonine, cysteine, glutamine and isoleucine. To the extent that replacement of asparagine with certain amino acids may give rise to interference with active site conformation, (e.g., due to steric hindrance which may be introduced by the presence of an aromatic amino acid or changes in tertiary structure such as may be introduced by the

- 7 -

presence of a proline) substitution with such amino acids would ordinarily be less preferred. Likewise, to the extent that replacement of asparagine with other amino acids may introduce a charged group (e.g.,

5 aspartic acid) into the proximity of the active site, such substitution would be less preferred. Illustrative of a presently preferred embodiment of an analog according to the present invention is a [Ser<sup>218</sup>]-analog of the aprA gene product. Alternative embodiments of 10 analogs within the contemplation of the invention are those wherein Asn<sup>109</sup> of subtilisin BPN' or of aprA gene product is replaced, preferably by a serine, and wherein glycine residues at positions 110 and/or 219 are replaced by different amino acid residues. In other 15 subtilisins, substitution for Asn at residue 62 or Gly at residue 63 of subtilisins Carlsberg or DY are also comprehended by the present invention.

A nucleic acid according to the present invention has codons encoding a polypeptide analog as 20 described above.

A system for the production of subtilisin according to the present invention comprises a host cell in turn comprising nucleic acid encoding an analog of subtilisin as described above. In such a cell, the 25 nucleic acid encoding the subtilisin analog may be chromosomal or extrachromosomal. The host cell is preferably selected from a strain deficient in secreted proteases, allowing for facile isolation of analog compounds.

30 A detergent formulation according to the present invention includes an analog of a Bacillus subtilisin having an amino acid sequence comprising an Asn-Gly sequence wherein one or both residues of the Asn-Gly sequence are deleted or are replaced.

35 A method for improving the thermal and pH stability of subtilisins according to the present inven-

- 8 -

tion comprises the step of substituting an amino acid other than asparagine for an asparagine in an Asn-Gly sequence and in particular for the asparagine residue at the position in the amino acid sequence of the 5 subtilisin which corresponds to position 218 in the amino acid sequence as disclosed in Table 1.

Brief Description of the Drawings

10 Fig. 1 schematically depicts the cyclization of Asn-Gly residues, such as those found at positions 218 and 219 of subtilisin as set forth in Table 1, to form anhydroaspartylglycine and also depicts base-catalyzed hydrolysis thereof;

15 Fig. 2 is a partial restriction map of an aprA gene-containing an EcoRI-KpnI gene fragment of Bacillus subtilis (B. subtilis) strain QB127 and includes a partial restriction map of the aprA gene and flanking sequences;

20 Fig. 3 is a partial restriction map of a plasmid pAMB11;

Fig. 4 is a flowchart depicting stages in construction of pAMB113, a plasmid which directs synthesis of [Ser]<sup>218</sup>-subtilisin from B. subtilis host 25 cells; and

Fig. 5 is a partial restriction map of pAMB30 plasmid.

Detailed Description

30

Bacillus serine proteases undergo irreversible inactivation in aqueous solutions at a rate that is largely dependent upon temperature and pH. At pH values below 4 or above 11 the rate of inactivation is very 35 rapid while at pH's of between 4.5 and 10.5 the rate, although much slower, increases as the solution becomes

- 9 -

more alkaline. In general, at any pH value, the higher the temperature the faster the rate of subtilisin deactivation.

A conserved sequence, Asn-Gly, at positions 5 109-110 and especially at positions 218-219 of Bacillus subtilisins is identified herein as a major factor responsible for the pH instability of these substances. The sequence Asn-Gly in proteins and peptides readily undergoes cyclization under various conditions, to form 10 the cyclic imide anhydroaspartylglycine [Bornstein, et al., Methods in Enzymol., 47, 132-145 (1977)], as illustrated in Fig. 1. This cyclic imide is susceptible to base-catalyzed hydrolysis which preferentially produces a non-native,  $\beta$ -aspartyl peptide bond. 15 Furthermore, the cyclic imide derived from Asn-Gly may serve as a target for specific cleavage of subtilisin. In fact, the specific cleavage of proteins at Asn-Gly bonds with alkaline hydroxylamine has become a common practice in the preparation of protein fragments for the 20 purpose of amino acid sequencing. [Bornstein, et al., supra]

Formation of a cyclic imide and/or  $\beta$ -aspartyl-glycyl peptide at Asn<sup>218</sup>-Gly<sup>219</sup> of subtilisin is predicted to cause irreversible inactivation of the 25 enzyme. This prediction is based on the close proximity of the unstable Asn-Gly element to a reactive serine located at position 221. A computer analysis of protein structures led to the belief that rearrangement of Asn<sup>218</sup>-Gly<sup>219</sup> to either anhydroaspartyl glycyl or to  $\beta$ -asparyl-glycyl results in a shift of the side-chain of 30 Ser<sup>221</sup> away from the position it must occupy for the enzyme to be active.

In order to eliminate the unstable element, Asn<sup>218</sup>-Gly<sup>219</sup>, from the subtilisin molecule one can 35 either replace Asn<sup>218</sup> with any amino acid other than asparagine and/or change Gly<sup>219</sup> to any amino acid other

- 10 -

than glycine. In a like manner, modification of the unstable Asn-Gly element at positions 109-110 is expected to provide benefits in the stability of analogs of the invention.

5        The observed invariance of Gly<sup>219</sup> in subtilisins and in subtilisin-like enzymes [e.g. Cucumisin from the melon Cucumis Melo L. Var Prince, Kaneda, et al., J. Biochem., 95, 825-829 (1984); and proteinase K, a subtilisin-like serine protease from the fungus

10      Tritirachium album, Jany et al., Biol. Chem. Hoppe-Seyler, 366, 485-492 (1985)] and the assumption that if Gly<sup>219</sup> were any residue other than glycine, its side chain might interfere with binding of a substrate to the enzyme, make highly preferred the change of Asn<sup>218</sup>

15      rather than Gly<sup>219</sup> for the removal of the unstable Asn<sup>218</sup>-Gly<sup>219</sup> sequence.

Based on theoretical considerations and on the compilation and analysis of sequencing data, asparagine at position 218 was replaced by serine in the presently preferred embodiment of the present invention described in the following illustrative examples. This selection was based in part on the observation that the amino acid sequence around the reactive serine of cucumisin, a subtilisin-like enzyme from melon fruit, has the sequence Ser-Gly-Thr-Ser-Met (Kaneda, et al., supra). Proteinase K has the same sequence around the active site. Jany et al., supra. It should be emphasized, however, that the selection of serine as a substitute for Asn<sup>218</sup> does not preclude achieving the same goal, i.e., elimination of the unstable element Asn<sup>218</sup>-Gly<sup>219</sup>, through replacing asparagine at position 218 with a different amino acid. It is preferred that an uncharged aliphatic amino acid, such as valine, threonine, cysteine, glutamine or isoleucine be substituted for Asn<sup>218</sup>.

Due to their capacity to secrete substantial

- 11 -

quantities of proteins and because they are currently used to produce detergent proteases, Bacillus micro-organisms represent a preferred host for recombinant production of the [Ser<sup>218</sup>]-subtilisin according to the 5 present invention. Because most Bacilli secrete alkaline and neutral proteases, it is preferable that mutations be introduced into the endogenous alkaline and neutral protease genes of B. subtilis so that the mutated subtilisin may be produced and secreted by B. 10 subtilis in a medium free of other proteases. Thus the present invention also provides mutant strains of B. subtilis which are blocked with respect to the synthesis of endogenous proteases but which retain the ability to synthesize and secrete subtilisin analogs such as 15 [Ser<sup>218</sup>]-subtilisin.

As described in greater detail below, it was found that the pH and thermal stability and the stability in detergent formulations of [Ser<sup>218</sup>]-aprA gene product subtilisin is much greater than that of the 20 wild type aprA gene product subtilisin.

The production of a stable subtilisin analog according to the invention included the following procedures:

1. Isolation of the representative subtilisin 25 gene apr A from B. subtilis;
2. Cloning of the apr A gene on a vector which permits utilization of oligonucleotide site-directed mutagenesis to create desired modifications;
3. Site-directed mutagenesis and sequencing 30 of the resulting DNA to confirm the presence of the desired mutation;
4. Construction of an expression vector to direct the synthesis of the mutated enzyme in B. subtilis;
- 35 5. Construction of mutated B. subtilis strains which do not synthesize subtilisin and neutral

- 12 -

protease;

6. Isolation of the enzyme in the extra-cellular growth medium and its purification;

7. Assessment of stability and activity

5 characteristics of the isolated product; and

8. Practice of procedures for insertion of the gene coding for the improved enzyme into the chromosome of a B. subtilis strain previously mutated to block synthesis of endogenous proteases.

10 In Example 1, the aprA gene encoding subtilisin is isolated from the B. subtilis genome. In Example 2, the aprA gene is subjected to site-directed mutagenesis. In Example 3, an expression vector containing the mutated aprA gene is constructed. In 15 Example 4, two mutant strains of B. subtilis which produce no detectable extracellular proteases are constructed. Example 5 describes procedures for integration of a mutated aprA gene into the chromosome of B. subtilis. In Example 6, wild-type and mutant aprA 20 subtilisins are isolated and purified. Examples 7 through 10 compare the thermostability of [Ser<sup>218</sup>] subtilisin to that of wild-type aprA gene product and to that of a commercial BPN' product.

25

Example 1

B. subtilis strain QB127 (trpC2 leuA8 30 sacU<sup>h</sup>200) [Lepesant, et al., Molec. Gen. Genet., 118, 135-160 (1982)] was obtained from the Bacillus Genetic Stock Center at the Ohio State University, Columbus, Ohio. This strain overproduces extracellular serine and metal proteases,  $\alpha$ -amylase and levansucrase relative to isogenic sacU<sup>+</sup> strains due to the pleiotropic effect of the sacU<sup>h</sup>200 mutation [Lepesant, et al., in 35 Schlessinger, D., ed., Microbiology, 1976, American Society for Microbiology, Washington, D.C., p. 65

- 13 -

(1976)] Strain QB127 was thus perceived to be a suitable source of DNA for isolating the aprA gene which codes for subtilisin.

Genomic DNA was isolated from cells of B. 5 subtilis strain QB127 by a published procedure [Saito, et al., Biochim. Biophys. Acta. 72, 619-629 (1963)]. Purified chromosomal DNA was digested to completion with the EcoRI restriction endonuclease.

DNA fragments were resolved on a low-melting 10 point agarose gel by electrophoresis and fragments in the 4.4 to 8.0 kilobase (kb) range were isolated by a standard DNA isolation procedure as recommended by the supplier (Bethesda Research Laboratories, Inc., Gaithersburg, Maryland). These fragments were ligated 15 to pCFM936 deposited as No. 53,413 with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland, on January 13, 1986, an Escherichia coli (E. coli) "walkaway" plasmid which displays higher copy numbers at elevated temperatures and which confers 20 kanamycin resistance.

The vector was digested with EcoRI and dephosphorylated with calf intestine alkaline phosphatase prior to ligation.

Ligation products were introduced into E. coli 25 C600 (available as A.T.C.C. 23724 from the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland) and kanamycin-resistant host cells were selected after overnight incubation on L-agar supplemented with 10 µg/ml kanamycin. Plasmid DNA was amplified by incubating host cells at 42°C for 4 hours. Colonies were then transferred to nitrocellulose filters and processed by a published procedure referred to as colony hybridization [Grunstein, et al., Proc. Natl. Acad. Sci. (USA), 72, 3961 (1975)].

35 A probe was used to screen for colonies which harbored the subtilisin gene on pCFM936. The probe

- 14 -

[synthesized by the phosphite chemistry method of Beau cage, et al., Tetrahedron Letters, 22, 1859-1862 (1981)] had the nucleotide sequence

(1) 5' GCGCAATCTGTTCCATTGGC 3'

5 which corresponds to the amino-terminus of the aprA gene product (Wong, et al., Proc. Natl. Acad. Sci. (USA), 81, 1184-1188 (1984); Stahl, et al., J. Bacteriol., 158, 411-418 (1984). A hybridization temperature of 55°C was used and 5 positive colonies were identified out of a 10 total of 400. The plasmid DNA from one of the positive colonies was designated pCFM936 apr2.

Plasmid pCFM936 apr2 was digested with EcoRI alone with HindIII alone and with EcoRI and HindIII in combination. Sizes of EcoRI fragments of the subtilisin 15 gene conformed to those described in Stahl, et al., supra, but several otherwise undescribed HindIII sites were discovered. As described in Example 3, two of the HindIII sites were employed for genetic manipulations of the subtilisin gene.

20 It was determined that a large 6.5 kb EcoRI fragment of B. subtilis QB127 genomic DNA carried the aprA gene, its regulatory sequences and unrelated flanking sequences by verifying that restriction enzyme digests conformed to the results reported by Stahl, et 25 al., supra. This was confirmed by DNA sequencing using the dideoxy chain termination method [Sanger, et al., J. Mol. Biol., 143, 161-178 (1980)]. A 3.0 kb EcoRI to KpnI subfragment of the 6.5 kb EcoRI fragment, as illustrated in Fig. 2, was also found to contain the 30 aprA gene, its regulatory sequences, and unrelated flanking sequences. Although the KpnI-EcoRI fragment is reported to be 2.5 kb in length in the text of Stahl, et al., and in the legend to Fig. 1 therein, comparison of the scale of Fig. 1 and the scaled depiction of the 35 fragment therein reveal that, even in Stahl, et al., the KpnI-EcoRI fragment is substantially larger than 2.5 kb.

- 15 -

A cloning vector for Bacillus host systems, plasmid pAMB11, was constructed as follows. The plasmid pTG402 (Northern Regional Research Laboratories, United States Department of Agriculture, Peoria, Illinois, 5 strain number NRRL B-15264) was partially digested with the RsaI restriction endonuclease. Fragments were ligated to M13 mpl8 (available from Bethesda Research Laboratories, Gaithersburg, Maryland as catalog number 8227SA) which had been previously digested with 10 HincII. Ligation products were introduced into E. coli JM103 (available from Pharmacia, Inc., Piscataway, New Jersey as catalog number 27-1545-01) by transformation [Mandel, et al., J. Mol. Biol., 53, 154, (1970)]. Bacteriophage plaques were sprayed with 0.5M catechol 15 (prepared in distilled water) to detect the functional expression of the xylE gene derived from pTG402. The xylE gene encodes catechol 2,3-dioxygenase and is useful for detecting promoters in a variety of organisms. Zukowski, et al., Proc. Natl. Acad. Sci. (USA), 80, 20 1101-1105 (1983).

The xylE gene was then transferred as a 1.0 kb EcoRI to PstI fragment to the E. coli/B. subtilis plasmid pHV33 (available from the American Type Culture Collection as A.T.C.C. 39217) [Primrose, et al. Plasmid, 25 6, 193-201 (1981)] obtained from R. Dedonder (Institut Pasteur, Paris, France). The pHV33 plasmid had been previously digested with EcoRI and PstI so that the xylE-containing fragment, when ligated in this region, would inactivate a gene for ampicillin resistance. The 30 resulting plasmid, pAMB21, contains a functional xylE gene in E. coli host cells, but requires the addition of a promoter for xylE to be expressed in B. subtilis host cells. E. coli cells harboring pAMB21 are resistant to tetracycline (15  $\mu$ g/ml) and chloramphenicol (20  $\mu$ g/ml) 35 while B. subtilis cells harboring pAMB21 are resistant only to chloramphenicol (5  $\mu$ g/ml).

- 16 -

The toop transcription termination sequence of bacteriophage lambda was then transferred from plasmid pCFM936 (on a 400 base pair PstI to BglII fragment) to the unique PstI site of pAMB21. A synthetic nucleotide 5 with the sequence 5' GATCTGCA 3' was constructed to join the BglII extremity of the toop fragment to the PstI site of the vector pAMB21. The resulting plasmid was designated pAMB22 and had properties indentical to pAMB21 except for the inclusion of a transcription 10 terminator. The pAMB22 plasmid is useful for detecting strong promoters that are functional in B. subtilis.

The 1.4 kb EcoRI to BglII fragment of DNA from pAMB22 that contains xylE and toop was isolated from a low-melting point agarose gel after electrophoresis of 15 restricted fragments. The 1.4 kb piece of DNA was ligated to plasmid pBD64 (available from Bacillus Genetic Stock Center, number 1E22) which had been previously digested with EcoRI and BamHI. The resulting 5.3 kb plasmid, pAMB11, contains the polylinker sequence 20 of M13mp18 (EcoRI, SstI, XmaI, Sma, BamHI and XbaI) upstream of the xylE gene which is followed by toop, as shown in Figure 3. The pAMB11 plasmid is capable of replicating in B. subtilis and confers upon host cells 25 resistance to chloramphenicol (5  $\mu$ g/ml) and/or kanamycin (5  $\mu$ g/ml).

As illustrated in Fig. 4, the purified EcoRI to KpnI fragment containing aprA was cloned onto pAMB11 to form pAMB111. Ligation products were introduced into 30 B. subtilis M1112 (arg-15 leuB thr5 recE4) (available from Bacillus Genetic Stock Center as No. 1A423) by the protoplast transformation method [Chang, et al., Mol. Gen. Genet., 168, 111-115 (1979)]. B. subtilis M1112 without plasmid DNA is protease-proficient (Prt<sup>+</sup> 35 phenotype), but secreted levels of subtilisin are rather low. Chloramphenicol-resistant (Cm<sup>R</sup>) transformants were transferred onto L-agar plates supplemented with 1.5%

- 17 -

(w/v) skim milk and 5  $\mu$ g/ml chloramphenicol, then incubated at 37°C.

After overnight (approx. 16 hrs.) incubation at 37°C, colonies of M112 harboring the new recombinant 5 plasmid (designated pAMB111) produced a clear halo surrounding each colony. Halos were formed by the proteolytic action of subtilisin on the casein component of the skim milk medium supplement. M112 harboring the pAMB111 vector alone had no visible halo after 16 hrs., 10 although a slight halo eventually developed after 40 hrs. at 37°C. Cells carrying pAMB111 were clearly distinguished from cells carrying pAMB11 by a difference in halo size. The cloning of the aprA gene in a fully 15 functional form was thus demonstrated to have led to the high level production and secretion of subtilisin by B. subtilis.

#### Example 2

20 As further illustrated in Fig. 4, the 3.0 kb EcoRI to KpnI genomic fragment, the isolation of which is described in Example 1, was digested with HindIII to produce three fragments: (1) a 1.1 kb EcoRI to HindIII fragment carrying genetic regulatory sequences for aprA 25 gene expression, the "pre-pro" region of the gene required to extracellular export of subtilisin, and the DNA sequence coding for the first 49 amino acids of mature subtilisin; (2) a 1.1 kb HindIII to HindIII fragment carrying DNA sequences coding for amino acids 50 through 275 (carboxyl-terminus) of subtilisin along with a transcription termination sequence and 3' non-coding 30 sequences; and (3) a 0.8 kb HindIII to KpnI fragment containing 3' non-coding sequences.

The 1.1 kb fragment flanked by HindIII sites 35 was cloned to the single HindIII site of bacteriophage M13 mp18 for the purposes of DNA sequencing and site-

- 18 -

directed mutagenesis. One of the recombinants, designated M13 mpl8 apr2, provided single stranded template DNA required for site-directed mutagenesis of the aprA gene.

5        The coding region of the aprA gene was sequenced and the results of the sequence are set out in Table 1 below. It should be noted that the specific identity of the initial 5 codons of the leader region is attributable to the report of Stahl, et al., supra, and

10 Wong, et al., supra, of sequence information for the aprA gene, and that there exist codon sequence differences from Stahl, et al., supra, at amino acid positions 84 and 85 which may be the result of sequencing error on the part of the authors of the

15 Stahl, et al. reference or which may be the result of a difference in the nucleotide sequences of the strains employed. Specifically, Stahl, et al., supra, reports a codon GTT (coding for valine) at amino acid position 84 while the codon GTA (also coding for valine) appears in

20 Table 1. Stahl, et al., supra, also reports a codon AGC (coding for serine) at amino acid position 85 as opposed to the codon GCG (coding for alanine) in Table 1.

25

30

35

- 19 -

TABLE 1

-105

fMet	Arg	Ser	Lys	Lys	Leu	Trp	Ile	Ser	Leu	Leu	Phe	Ala
GTG	AGA	AGC	AAA	AAA	TTG	TGG	ATC	AGC	TTG	TTG	TTT	GCG

Leu Thr Leu Ile Phe Thr Met Ala Phe Ser Asn Met Ser Ala  
TTA ACG TTA ATC TTT ACG ATG GCG TTC AGC AAC ATG TCT GCG

Gln Ala Ala Gly Lys Ser Ser Thr Glu Lys Lys Tyr Ile Val  
CAG GCT GCC GGA AAA AGC AGT ACA GAA AAG AAA TAC ATT GTC

```

Gly Phe Lys Gln Thr Met Ser Ala Met Ser Ser Ala Lys Lys
GGA TTT AAA CAG ACA ATG AGT GCC ATG ACT TCC CCC AAC AAA

```

Lys Asp Val Ile Ser Glu Lys Gly Gly Lys Val Gln Lys Gln  
AAG GAT GTT ATT TCT GAA AAA GGG GGA AAG GTT CAA AAC CAA

```

Phe Lys Tyr Val Asn Ala Ala Ala Ala Thr Leu Asp Glu Lys
TTT AAG TAT GTT AAC GCG GCG GCA GCA ACA TTG CAT CAA AAA

```

Ala Val Lys Glu Leu Lys Lys Asp Pro Ser Val Ala Tyr Val  
GCT GTA AAA GAA TTG AAA AAA GAT CGG AGC GTC GCA TAT CTC

	-1    +1													
Glu   Glu   Asp   His   Ile   Ala   His   Glu   Tyr   Ala   Gln   Ser   Val   Pro	GAA	GAA	GAT	CAT	ATT	GCA	CAT	GAA	TAT	GGC	CAA	TCT	CTT	CCT

10  
Tyr Gly Ile Ser Gln Ile Lys Ala Pro Ala Leu His Ser Gln  
TAT GGC ATT TCT CAA ATT AAA GCG CCG GCT CTT CAC TCT CAA

20	30
Gly Tyr Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp Ser	GGC TAC ACA GGC TCT AAC GTC AAA GTC CCT CTT ATC GAC AGC

40

50	60
Ala Ser Phe Val Pro Ser Glu Thr Asn Pro Tyr Gln Asp Gly	CCA ACC TTG CTA CCT TGT GAA AGC AGG TGT GTC GAA GCG

70

Ser	Ser	His	Gly	Thr	His	Val	Ala	Gly	Thr	Ile	Ala	Ala	Leu
AGT	TCT	CAC	GCT	ACG	CAT	CTA	CCC	CCT	AGC	ATM	CCC	CCC	CTT

- 20 -

TABLE 1 (cont'd.)

80  
 Asn Asn Ser Ile Gly Val Leu Gly Val Ala Pro Ser Ala Ser  
 AAT AAC TCA ATC GGT GTT CTG GGC GTA GCG CCA AGC GCA TCA

90 100  
 Leu Tyr Ala Val Lys Val Leu Asp Ser Thr Gly Ser Gly Gln  
 TTA TAT GCA GTA AAA GTG CTT GAT TCA ACA GGA AGC GGC CAA

110  
 Tyr Ser Trp Ile Ile Asn Gly Ile Glu Trp Ala Ile Ser Asn  
 TAT AGC TGG ATT ATT AAC GGC ATT GAG TGG GCC ATT TCC AAC

120 130  
 Asn Met Asp Val Ile Asn Met Ser Leu Gly Gly Pro Thr Gly  
 AAT ATG GAT GTT ATC AAC ATG AGC CTT GGC GGA CCT ACT GGT

140  
 Ser Thr Ala Leu Lys Thr Val Val Asp Lys Ala Val Ser Ser  
 TCT ACA GCG CTG AAA ACA GTC GTT GAC AAA GCC GTT TCC AGC

150  
 Gly Ile Val Val Ala Ala Ala Ala Gly Asn Glu Gly Ser Ser  
 GGT ATC GTC GTT GCT GCC GCA GCC GGA AAC GAA GGT TCA TCC

160 170  
 Gly Ser Thr Ser Thr Val Gly Tyr Pro Ala Lys Tyr Pro Ser  
 GGA AGC ACA AGC GTC GGC TAC CCT GCA AAA TAT CCT TCT

180  
 Thr Ile Ala Val Gly Ala Val Asn Ser Ser Asn Gln Arg Ala  
 ACT ATT GCA GTA GGT GCG GTA AAC AGC AGC AAC CAA AGA GCT

190 200  
 Ser Phe Ser Ser Ala Gly Ser Glu Leu Asp Val Met Ala Pro  
 TCA TTC TCC AGC GCA GGT TCT GAG CTT GAT GTG ATG GCT CCT

210  
 Gly Val Ser Ile Gln Ser Thr Leu Pro Gly Gly Thr Tyr Gly  
 GGC GTG TCC ATC CAA AGC ACA CTT CCT GGA GGC ACT TAC GGC

220  
 Ala Tyr Asn Gly Thr Ser Met Ala Thr Pro His Val Ala Gly  
 GCT TAT AAC GGA ACG TCC ATG GCG ACT CCT CAC GTT GCC GGA

230 240  
 Ala Ala Ala Leu Ile Leu Ser Lys His Pro Thr Trp Thr Asn  
 GCA GCA GCG TTA ATT CTT TCT AAG CAC CCG ACT TGG ACA AAC

250  
 Ala Gln Val Arg Asp Arg Leu Glu Ser Thr Ala Thr Tyr Leu  
 GCG CAA GTC CGT GAT CGT TTA GAA AGC ACT GCA ACA TAT CTT

- 21 -

TABLE 1 (cont'd.)

260 270  
Gly Asn Ser Phe Tyr Tyr Gly Lys Gly Leu Ile Asn Val Gln  
GGA AAC TCT TTC TAC TAT GGA AAA GGG TTA ATC AAC GTA CAA

275  
Ala Ala Ala Gln OC  
GCA GCT GCA CAA TAA TAGTAAAAAGAAGCAGGTTCCCTCCATACCTGCT  
TCTTTTATTGTCAGCATTGATGTTCCGGCGCATTCTC

- 22 -

Site-directed mutagenesis was performed by a standard method. Norrander, et al., Gene, 26, 101-106 (1983). Single-stranded DNA from M13 mpl8 apr2 was 5 annealed to the mutagenic primer

(2) 5' GGCGCTTATAGCGGAAC 3'

which was synthesized by the phosphite chemistry method (Beaucage, et al., supra). The synthetic primer was homologous to codons for amino acids 216 through 220 of 10 subtilisin with the exception of a single base change in the codon for amino acid 218 (AGC instead of AAC). Such a change allowed the mutagenic reaction to substitute a serine codon at position 218 in place of the original asparagine codon at this position.

15 The synthetic primer was annealed to M13 mpl8 apr2 DNA at 65°C before being slowly cooled to room temperature (approx. 22°C). Polymerization followed for 2 hr. at 15°C in a reaction mixture which consisted of 12.5 µl of annealed DNA solution, 2.5 µl of 10 mM each 20 dATP, dGTP, dCTP and dTTP, 2.0 µl of 12 mM ATP, 0.1 µl Klenow DNA polymerase, 0.1 µl T4 DNA ligase and 13 µl sterile distilled water. The resulting double-stranded, covalently closed circular DNA was introduced into E. coli JM103 by transfection.

25 Bacteriophage plaques were then transferred to Gene Screen™ (New England Nuclear, Beverley, Massachusetts) hybridization membranes. Plaques which contained DNA with the desired base change were identified by hybridization to radioactively labeled ( $\gamma$ -<sup>32</sup>P) 30 synthetic oligonucleotide (2) used for the mutagenic priming reaction described above. Hybridization was performed at a restrictive temperature (52°C) so that only DNA carrying the Ser<sup>218</sup> mutation would hybridize to the synthetic oligonucleotide. The presence of the 35 Ser<sup>218</sup> mutation in the aprA gene on DNA from a single purified plaque, designated M13 mpl8 apr2 [Ser]<sup>218</sup>, was

- 23 -

confirmed by DNA sequencing by the method of Sanger, et al., supra.

Example 3

5

In order to express [Ser<sup>218</sup>]-subtilisin in B. subtilis, a suitable plasmid vehicle was constructed by digesting pAMB111 with HindIII. The 1.1 kb segment carrying most of the aprA gene was deleted by re-  
10 ligating HindIII digestion products of pAMB111 at a concentration of approximately 1 $\mu$ g/ml. This resulted in the formation of pAMB110 as illustrated in Fig. 4. The pAMB110 plasmid carries genetic regulatory sequences for expression of the subtilisin gene, the "pre-pro" region  
15 required for secretion of subtilisin, and the DNA sequence coding for the 3' non-coding region of mature subtilisin and the first 49 amino acids of mature subtilisin. Because it is lacking DNA coding for amino acids 50 through 275, pAMB110 does not synthesize sub-  
20 tilisin when introduced into B. subtilis host cells. Subtilisin is synthesized only after insertion of the remainder of a subtilisin gene, either the native DNA sequence or an analog-encoding sequence, such as a sequence encoding [Ser<sup>218</sup>]-subtilisin.  
25 Double-stranded DNA from M13 mpl8 apr2 [Ser]<sup>218</sup> was digested with HindIII. A 1.1 kb fragment carrying the aprA gene segment with the Ser<sup>218</sup> mutation was then ligated to pAMB110 which had been previously digested with HindIII. Ligation products were intro-  
30 duced into B. subtilis by transformation as in Example 1 above. Ligation of the 1.1 kb HindIII fragment in the correct orientation (as confirmed by DNA sequencing by the method of Sanger, et al. supra) for expression of the mutated gene resulted in the construction of  
35 pAMB113, a plasmid which directed synthesis and secretion of [Ser<sup>218</sup>]-subtilisin from B. subtilis host cells.

- 24 -

Example 4

Because most Bacilli secrete alkaline and/or neutral proteases into the surrounding growth medium, it is preferable that mutations be introduced into endogenous alkaline and neutral protease genes of B. subtilis to block their synthesis so that mutated subtilisin genes, when introduced into the mutant cell, may produce mutated subtilisins which will then be secreted in a medium free of other proteases likely to interfere with isolation of intact subtilisin analogs. Two mutant B. subtilis strains BZ24 and BZ25, which produce no detectable extracellular proteases, were constructed in the following manner.

First, a plasmid vehicle capable of replicating in E. coli, but not in B. subtilis unless integrated into the B. subtilis chromosome by homologous recombination, was constructed as follows. Plasmid pBD64 (Bacillus Genetic Stock Center, Number 1E22) was digested to completion with HpaII to produce three fragments of 2.95 kb, 1.0 kb and 0.75 kb in size. These fragments were then ligated as a mixture to plasmid pBR322 (A.T.C.C. 37017) which previously had been digested with ClaI. Ligation products were introduced into E. coli C600 (available from the American Type Culture Collection as A.T.C.C. 23724) by transformation [Mandel, et al., J. Mol. Biol., 53, 154 (1970)]. Selection was for cells resistant to chloramphenicol (20  $\mu$ g/ml) and ampicillin (50  $\mu$ g/ml). Plasmid DNA from 12 transformants was prepared by an alkaline extraction procedure [Birnboim, et al., Nucleic Acids Res., 7, 1513-1523 (1979), then digested with HindIII and EcoRI in combination to verify the presence of inserted fragment(s). One such plasmid, designated pAMB30, was found to carry the 1.0 and 0.75 kb HpaII fragments of pBD64 in the ClaI site of pBR322. These fragments contain the

- 25 -

chloramphenicol acetyltransferase (cat) gene which is functional in E. coli and B. subtilis. Digestions with BglIII and, separately, with Sau3A confirmed the identity and orientation of the cat gene on pAMB30, as 5 illustrated in Fig. 5.

Because pAMB30 lacks an origin of replication sequence which is functional in B. subtilis, it cannot replicate as an autonomous replicon in B. subtilis host cells. On the other hand, pAMB30 contains the pBR322-10 derived origin of replication which is functional in E. coli, thus the plasmid can be propagated in E. coli host cells. Plasmid pAMB30 is useful in at least 2 ways. First, a fragment of DNA which contains a functional origin of replication in B. subtilis may be detected 15 when cloned onto pAMB30 such that the plasmid will autonomously replicate in the extrachromosomal state. Second, plasmid pAMB30 can integrate into the genome of B. subtilis at a site of homology between the chromosome and B. subtilis DNA cloned onto pAMB30. Such an event 20 has been repeatedly demonstrated in the past [Haldenwang, et al., J. Bacteriol., 142, 90-98 (1980); Young, J. Gen. Microbiol., 129, 1497-1512 (1983)] by using plasmid vehicles similar to, but not identical with, pAMB30.

25 Plasmid pAMB21 (described in Example 1) was digested with EcoRI and PstI to isolate the xylE gene on a 1.0 kb fragment. The fragment was ligated to pAMB30 which had been previously digested with EcoRI and PstI. Ligation products were introduced into E. coli 30 C600 by transformation. Selection was for chloramphenicol resistant (20  $\mu$ g/ml) host cells which were sensitive to ampicillin (50  $\mu$ g/ml) due to the insertion of the xylE fragment of pAMB21 into the structural gene for ampicillin resistance of pAMB30. The resulting 35 plasmid, pAMB30/21, has properties identical to pAMB30 but has, in addition, a functional xylE gene.

- 26 -

Plasmid pAMB110, which carries the aprA gene deleted of a region coding for the latter 226 amino acids of mature subtilisin, was digested with EcoRI and KpnI. The 1.9 kb fragment of B. subtilis DNA containing 5 genetic regulatory sequences for aprA gene expression, "the pre-pro" region, the DNA sequence coding for the first 49 amino acids of mature subtilisin and 3' non-coding sequences was ligated to pAMB30/21 that had been previously digested with EcoRI and KpnI. Ligation 10 products were introduced into E. coli C600 by transformation. Plasmid DNA from several transformants was isolated by the alkaline extraction procedure of Birnboim, et al., supra, and the presence of the inserted 1.9 kb fragment was verified by multiple 15 restriction endonuclease digestions. One such plasmid, designated pAMB301, was retained for further use.

B. subtilis strain BGSC1A274 (Bacillus Genetic Stock Center) carries a mutation at the npr locus and is incapable of producing extracellular neutral protease. 20 The plasmid pAMB301 was integrated into the genome of B. subtilis BGSC1A274 by transformation of competent cells [Spizizen, Proc. Natl. Acad. Sci. (USA), 44, 1072-1078 (1958)]. Selection was for chloramphenicol-resistant (5  $\mu$ g/ml) host cells which were then transferred by sterile 25 toothpicks to L-agar supplemented with 1.5% (w/v) powdered skim milk and (5  $\mu$ /ml) cloramphenicol. Those cells which failed to produce a clear halo surrounding the colony were deficient in the ability to produce extracellular neutral and serine proteases due to the 30 combination of the npr mutation along with the newly introduced aprA mutation. The aprA mutation was a deletion of the latter 226 amino acids of mature subtilisin due to the replacement of the wild-type aprA gene with the deleted version carried on pAMB301. One 35 such strain, designated BZ24, has the Npr<sup>-</sup> Apr<sup>-</sup> Cm<sup>r</sup> phenotype, thus it produces no detectable extracellular

- 27 -

neutral protease nor extracellular alkaline protease and is resistant to chloramphenicol at 5  $\mu$ g/ml. Southern blotting [Southern, J. Mol. Biol., 98, 503-517 (1975)] was used to confirm the deletion in the aprA gene on the 5 chromosome of B. subtilis BZ24. Cultivation of B. subtilis BZ24 in Antibiotic Medium No. 3 (Penassay Broth, Difco, Detroit, Michigan) in the absence of antibiotic selection for approximately 32 generations led to the isolation of a derivative strain of BZ24 in 10 which the cat gene conferring chloramphenicol resistance upon host cells was lost due to its instability in the BZ24 chromosome. Such a phenomenon has been previously observed in similar experiments [Stahl, et al., J. Bacteriol., 158, 411-418 (1984)]. A chloramphenicol-sensitive derivative of BZ24 was designated BZ25. B. subtilis BZ25 has the Npr<sup>-</sup> Apr<sup>-</sup> phenotype, thus it produces no detectable extracellular neutral protease nor extracellular alkaline protease. Southern blotting was used to confirm the deletion in the aprA gene on the 15 20 chromosome of B. subtilis BZ25.

Because B. subtilis BZ25 produces no detectable extracellular neutral protease nor subtilisin, it is a useful host strain for introduction of plasmid DNA, such as pAMB113, for the production of 25 mutated subtilisins which may be secreted into the surrounding growth medium free of other proteases.

B. subtilis BZ25 produces no detectable extracellular proteases when culture supernatants are assayed as described below. B. subtilis BZ25/pAMB113, which is 30 BZ25 that harbors plasmid pAMB113 (introduced by the protoplast transformation method of Chang, et al., supra) produces appreciable quantities of [ $\text{Ser}^{218}$ ]-subtilisin when culture supernatants are assayed as described.

- 28 -

Example 5

Integration of the [Ser<sup>218</sup>]-subtilisin gene into the chromosome of B. subtilis was believed to 5 provide an efficient way of increasing genetic stability of this mutant gene. Such an approach also alleviates the requirement for chloramphenicol in the fermentation medium which is otherwise needed for application of selective pressure to maintain plasmid DNA in the extra- 10 chromosomal state. Therefore, the [Ser<sup>218</sup>]-subtilisin gene, along with its genetic regulatory sequences and flanking DNA homologous to the B. subtilis chromosome, was isolated from a low melting point agarose gel after electrophoresis of pAMB113 which had been digested with 15 EcoRI and PstI in combination. The 4.0 kb EcoRI to PstI fragment (illustrated in Fig. 4) was then ligated to pAMB30 (illustrated in Fig. 5) which had been digested with EcoRI and PstI in combination. Ligation products were introduced into E. coli HB101 (A.T.C.C. 33694) by 20 transformation. Selection was for cells resistant to chloramphenicol (20 µg/ml). Plasmid DNA from four transformants which met the criteria above were isolated by the alkaline extraction procedure of Birnboim, et al., supra, then digested with EcoRI and PstI in 25 combination. All four plasmids contained the 4.0 kb insert and the 5.6 kb remaining portion of pAMB30. One such plasmid, designated pAMB302, was purified and retained for further use.

Repeated attempts to integrate plasmid pAMB302 30 into the chromosome of B. subtilis BZ25 by the competence method [Spizizen, supra] were unsuccessful. This may have been due to the failure of BZ25 cells to become competent by the method employed. Therefore, pAMB302 was introduced into B. subtilis BZ25 cells by 35 the protoplast transformation method of Chang, et al., supra. This is believed to be the first demonstration

- 29 -

that the protoplast transformation method is successful for obtaining integration of heterologous DNA in Bacillus. This result is particularly significant in that research strains in which integration has been 5 obtained were selected on the basis of transformation by the competence method. Strains which may be unable to become competent, and in particular industrial strains which were not selected on the basis of transformation by the competence method, may be more likely to be 10 unable to become competent.

Selection was for chloramphenicol-resistant cells (5  $\mu$ g/ml) cells, which were then transferred with sterile toothpicks to L-agar supplemented with 1.5% (w/v) skim milk and 5  $\mu$ g/ml chloramphenicol. Cells were 15 incubated overnight at 37°C. Clear halos of different diameters were observed around the  $Cm^r$  colonies. This indicates that subtilisin was produced and secreted by these cells. An attempt was made to isolate plasmid DNA from eight of these colonies by the alkaline extraction 20 method. No plasmid DNA was detected on agarose gels which were stained with ethidium bromide (1  $\mu$ g/ml) to visualize DNA after electrophoresis. The absence of extrachromosomal plasmid DNA in the  $Cm^r$  cells which produced subtilisin was a strong indication that pAMB302 25 had been integrated into the chromosome of B. subtilis.

Several colonies resulting from this experiment were isolated and designated BZ28, BZ29, BZ30, BZ31, BZ32 and BZ33. Each strain was grown 30 overnight at 37°C with vigorous shaking in brain heart infusion medium (BHI, Difco) supplemented with 5  $\mu$ g/ml chloramphenicol. Culture supernatants were assayed for subtilisin activity. B. subtilis strains BZ28, BZ29, BZ30, BZ31, BZ32 and BZ33 all produced subtilisin and 35 secreted it into the surrounding growth medium, some strains producing more than others. The amount of sub-

- 30 -

tilisin observed in the liquid culture broth was directly proportional to the size of the halo observed on skim milk L-agar plates. Because of the amounts of subtilisin secreted by these cells differed, it was 5 postulated that either multiple copies of pAMB302 were integrated into the chromosome or that gene amplification [Young, J. Gen. Microbiol., 129, 1497-1512 (1983); Albertini, et al., J. Bacteriol., 162, 1203-1211 (1985)] had taken place.

10

Example 5

Wild-type subtilisin, from BZ25/pAMB111, and [Ser<sup>218</sup>]-subtilisin, from BZ25/pAMB113, were isolated 15 and purified as follows. Each culture broth was centrifuged at 15,000g for 30 minutes and protein in the clear supernatant was precipitated with (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (350g per liter). The precipitate was collected by centrifugation and after trituration with 75% acetone, 20 it was filtered and dried under a vacuum.

In order to further purify the enzyme, the dried precipitate was dissolved in water and the solution was filtered and then dialyzed against 0.02M sodium phosphate buffer at pH 6.3. The dialyzed solution was 25 passed through a column (2.5 x. 15cm) of carboxymethyl cellulose at a rate of 2 ml per minute. After washing the column with 0.02M sodium phosphate (pH 6.3), the enzyme was eluted with the same buffer containing 0.15M NaCl. Peak fractions were pooled and protein from the 30 fractions containing the enzyme, as identified by a color change in a sample of the fraction mixed with succinyl-L-alanyl-L-alanyl-L-prolyl-L-phenylalanyl-p-nitroanilide (Vega Biochemicals), were precipitated by addition of 2.5 volumes of acetone. The precipitate was 35 collected by centrifugation and then dissolved in 0.005M calcium acetate (about 1 ml per 10 mg). The resulting

- 31 -

solution was dialyzed at 4°C against water and then lyophilized.

Example 7

5        Because the present invention relates to the stabilization of Bacillus subtilisin, an analysis of enzyme stability and quantitation thereof was undertaken.

10      Most enzymes express their biological activity, i.e., catalysis of chemical reactions, only within a narrow range of pH and temperature. Moreover, even under optimal conditions, the enzyme will retain its activity only if its polypeptide chain is folded in a manner that forms the so-called native conformation.

15      The native form of an enzyme is thermodynamically more stable than the denatured or unfolded form by an average of 10-15 kilocalories per mole. Unfolding of the native structure often occurs when the enzyme is exposed to extremes of pH or temperature, or to certain concentrations of chemicals such as detergents, urea and organic solvents. Removal of these denaturing agents often result in spontaneous refolding of the peptide chain to the native form, and in restoration of the original enzyme activity.

25      Irreversible loss of enzyme activity may occur due to cleavage of the polypeptide chain or due to modification of certain amino acid side chains, especially if these modifications alter the native architecture of the enzymes' active site. Examples of 30 such modifications include the deamidation of asparaginyl and glutaminyl residues, the oxidation of methionyl residues and the hydrolytic cleavage of cysteine to form one residue of thiocysteine and one of dehydroalanine. The present invention provides an 35 additional example in the form of the irreversible inactivation through cyclization of Asn-Gly sequences.

- 32 -

When the enzyme preparation includes several enzymatic forms which inactivate at different rates and/or when the inactivation process occur through a number of mechanisms, the kinetics of inactivation are 5 complicated. However, for most enzyme preparations at a suitable range of pH and temperature, the thermal inactivation follow first-order kinetics, i.e., the residual enzyme activity decreases as a function of time along an exponential decay curve. Under these 10 conditions the half-life ( $T_{1/2}$ ) of the enzyme is independent of the initial enzyme concentration and may be calculated according to:

$$15 \quad T_{1/2} = \frac{(t_2 - t_1) \ln 2}{\ln A_1 - \ln A_2}$$

in which  $A_1$  and  $A_2$  are the enzyme activities at times  $t_1$  and  $t_2$  respectively.

20 In general, everything else being equal, the half-life of an enzyme in solution is shorter at higher temperatures.

In order to compare the thermostability of the 25  $[Ser^{218}]\text{-}\underline{aprA}$  gene product subtilisin to that of wild-type aprA gene product subtilisin and subtilisin BPN' (Sigma), solutions of these enzymes (1 mg/ml) were prepared in 0.1M sodium glycinate buffer at pH 10.0. The solutions were incubated at 52°C and after various 30 times aliquots (20  $\mu$ l) were drawn and mixed with 900  $\mu$ l of 0.2% casein solution in 0.1M Tris buffer at pH 8.30. As a control, the substrate (casein) solution was incubated with 20  $\mu$ l of enzyme buffer. The hydrolysis 35 of casein at room temperature was terminated after 15 minutes with the addition of 200  $\mu$ l of 10% trichloroacetic acid. The hydrolysate was separated from the precipitated protein by centrifugation and its

- 33 -

ultraviolet (UV) absorbance at 280nm, as compared to the control, was measured by an 8451A Diode Array Spectrophotometer available from Hewlett Packard. Substrate concentration was such that enzyme activities 5 were directly proportional to the UV absorbances at 280nm reported in Table 2. In this table the values in parentheses represent percentages of initial enzyme activities. The last column in Table 2 shows the calculated half-life of the three enzymes, and it can be 10 seen that the half-life of the mutated  $[\text{Ser}^{218}]_{\text{aprA}}$  gene product is more than three-fold longer than the natural aprA gene product and subtilisin BPN' under the tested conditions.

15

20

25

30

35

- 34 -

**TABLE 2**  
**Thermostability of  $[Ser218]^{aprA}$  Gene Product vs. Wild-Type  $aprA$  Gene Product**  
**and Subtilisin BPN' in 0.1M Sodium Glycinate at pH 10.0**

<u>Protease</u>	Enzyme Activity After Incubation At 52°C For:			<u>Half-Life of the Enzyme</u>
	<u>0. Hour</u>	<u>1 Hour</u>	<u>2 Hour</u>	
$[Ser218]^{aprA}$	0.993 (100%)	0.887 (89%)	0.738 (74%)	0.620 (62%)
Wild-Type $aprA$	0.132 (100%)	0.702 (62%)	0.437 (38%)	0.252 (22%)
Subtilisin BPN'	0.998 (100%)	0.596 (60%)	0.360 (36%)	0.192 (19%)

- 35 -

Example 8

For determination of thermostability of subtilisins in the presence of detergents, the liquid 5 laundry detergent ERA Plus® (Procter & Gamble) was used after diluting it with water (1:9) and completely inactivating the original protease activity by warming it to 65°C for 30 minutes. The pH of the resultant detergent solution was 7.50. Using the casein assay and the 10 procedure described in Example 7, the stabilities of [Ser<sup>218</sup>]aprA gene product, wild-type aprA gene product and subtilisin BPN' (Sigma) were tested in the detergent at 45°C. The results are shown in Table 3 in which 15 enzyme activities are expressed as percentages of original enzyme activities. Again, the half-life of the subtilisin analog was on the order of three-fold greater than the natural products under the test conditions.

20

25

30

35

TABLE 3

Thermostability of  $[{\text{Ser}}^{218}]aprA$  Gene Product vs. Wild-Type  $aprA$  Gene Product and Subtilisin BPN' in Detergent at pH 7.5

<u>Protease</u>	<u>Enzyme Activity After Incubation At 45°C For:</u>			<u>Half-Life of the Enzyme</u>
	<u>0 Hour</u>	<u>0.5 Hour</u>	<u>1.5 Hour</u>	
$[{\text{Ser}}^{218}]aprA$	100%	93%	82%	6.0 hr
Wild-Type $aprA$	100%	80%	54%	1.73 hr
Subtilisin BPN'	100%	84%	62%	2.15 hr

- 37 -

The above-tabulated results demonstrate that the [Ser<sup>218</sup>]-analog of the aprA gene product exhibits greater stability than the wild type aprA gene product or subtilisin BPN' in ERA Plus® which, with a pH of 7.5, 5 may be described as a cationic to neutral detergent and which most likely has been formulated to be compatible with inclusion of detergent enzymes. These results do not assure, however, that the [Ser<sup>218</sup>]-aprA gene product will be compatible with all current commercial detergent 10 formulations, e.g., those which have been formulated to the exclusion of detergent enzymes. For example, in preliminary tests with a 2% (w/v) solution of Tide®, which may be described as an anionic to neutral detergent having a pH of greater than 8.5 and which 15 excludes detergent enzymes in its formulation, the subtilisin BPN' exhibited greater stability than the wild-type and [Ser<sup>218</sup>]-aprA gene products. The [Ser<sup>218</sup>]-aprA gene product in that same test, however, showed greater stability than the wild-type aprA gene 20 product. Although the different relative performance of the [Ser<sup>218</sup>]-aprA gene product and subtilisin BPN' in ERA Plus® and Tide® is as yet unexplained, the experimental results suggest that proper formulation of detergent compositions is a prerequisite for optimal 25 performance of enzymes included in such compositions. What has been clearly demonstrated is that the [Ser<sup>218</sup>]-aprA gene product consistently possesses properties superior to those of the wild type aprA gene product, and it is believed that analogs of subtilisins Carlsberg 30 and BPN' according to the present invention will also possess greater stability than the corresponding wild type enzyme.

- 38 -

Example 9

Using succinyl-L-alanyl-L-alanyl-L-prolyl-L-phenylalanyl-p-nitroanilide (Vega Biochemicals) as 5 substrate and the rate of increase in absorbance at 405nm due to release of p-nitroaniline [Del Mar, et al., Anal. Biochem., 99, 316-320, (1979)] to measure enzyme activity, the thermostabilities of [ $\text{Ser}^{218}$ ]aprA, wild-type aprA and subtilisin BPN' (Sigma) in 0.1M sodium 10 phosphate buffer at pH 7.5 were determined as follows.

Enzyme solutions of about 0.5 Anson units per liter were incubated at 40°C and 50°C, and at various times aliquots (20  $\mu$ l) were drawn and diluted into 180  $\mu$ l of ice-cold 0.1M sodium phosphate buffer at pH 7.5. 15 10  $\mu$ l of the thus diluted sample was mixed with 890  $\mu$ l of 1mM succinyl-L-alanyl-L-alanyl-L-prolyl-L-phenylalanyl-p-nitroanilide in 0.1M Tris HCl pH 8.2 and the absorbances at 405 nm were measured every 15 seconds for 5 minutes on a Hewlett Packard, 8451A Diode Array 20 Spectrophotometer. Residual enzyme activities after various incubation times are expressed in Table 4 and Table 5 as percentages of the corresponding initial activities. The procedure was then repeated in 0.1M sodium phosphate buffer at pH 9.0 and the results are 25 set out in Tables 6 and 7.

- 39 -

TABLE 4

Thermostability of  $[{\text{Ser}}^{218}]^{\text{aprA}}$  Gene Product vs. Wild-Type  $\underline{\text{aprA}}$  Gene Product and Subtilisin BPN' in 0.1M Sodium Phosphate Buffer at pH 7.5

Protease	Enzyme Activity After Incubation at 40°C For:			Half Life of the enzyme
	0. Min.	90 Min.	180 Min.	
$[{\text{Ser}}^{218}]^{\text{aprA}}$	100%	92%	87%	13.0 hr
Wild-Type $\underline{\text{aprA}}$	100%	69%	48%	2.8 hr
Subtilisin BPN'	100%	48%	26%	1.60 hr
			14%	

- 40 -

TABLE 5

Thermostability of [Ser 218] aprA Gene Product vs. Wild-Type aprA Gene Product  
and Subtilisin BPN' in 0.1M Sodium Phosphate Buffer at pH 7.5

<u>Protease</u>	Enzyme Activity After Incubation At 50°C For:		<u>135 Min.</u>	<u>Half Life of the enzyme</u>
	<u>0 Min.</u>	<u>45 Min.</u>		
[Ser 218] <u>aprA</u>	100%	89%	81%	71%
Wild-Type <u>aprA</u>	100%	59%	35%	21%
Subtilisin BPN'	100%	34%	11%	4%

- 41 -

TABLE 6

Thermostability of  $[{\text{Ser}}^{218}]^{\text{aprA}}$  Gene Product vs. Wild-Type  $\text{aprA}$  Gene Product and Subtilisin BPN' in 0.1M Sodium Phosphate Buffer at pH 9.0

<u>Protease</u>	<u>Enzyme Activity After Incubation At 40°C For:</u>		<u>Half Life of the enzyme</u>
	<u>0 Min.</u>	<u>45 Min.</u>	
$[{\text{Ser}}^{218}]^{\text{aprA}}$	100%	96%	87%
Wild-Type $\text{aprA}$	100%	79%	49%
Subtilisin BPN'	100%	68%	29%

TABLE 7

Thermostability of [Ser 218]aprA Gene Product vs. Wild-Type aprA Gene Product and Subtilisin BPN' in 0.1M Sodium Phosphate Buffer at pH 9.0

Protease	Enzyme Activity After Incubation At 50°C For:			Half Life of the enzyme
	0 Min.	20 Min.	40 Min.	
[Ser 218] <u>aprA</u>	100%	92%	85%	69% 2.5 hr
Wild-Type <u>aprA</u>	100%	61%	36%	14% 0.47 hr
Subtilisin BPN'	100%	54%	31%	17% 0.37 hr
				9%

- 43 -

Example 10

In order to test pH stability of the [Ser<sup>218</sup>] analog, the procedure of Example 7 was carried out at pH 5 4.8 and room temperature and the results are set out in Table 8.

TABLE 8

10

Stability of [Ser<sup>218</sup>]aprA Gene Product vs.  
Wild-Type aprA Gene Product in 0.1M Sodium  
Phosphate Buffer at pH 4.8

15	<u>Protease</u>	<u>0 Hours</u>	<u>4 Days</u>	<u>Half-life of the enzyme (days)</u>
	[Ser <sup>218</sup> ] <u>aprA</u>	100	95%	44.0
	Wild-Type <u>aprA</u>	100	47%	3.7

20

While the present invention has been described in terms of preferred embodiments it is understood that modifications and improvements will occur to those skilled in the art. For example, the sequence Asn-Gly 25 appears at other points in subtilisins, such as at residues 109 and 110 of the aprA gene product and subtilisin BPN' and at residues 62 and 63 of subtilisin Carlsberg and of subtilisin DY. Thus, it is expected that substitution of residues other than Asn and Gly at 30 these respective locations may improve stability as well. Similar improvements in stability are expected for such substitutions made in other enzymes which have the Asn-Gly sequence and in other proteins comprising this sequence. Furthermore, it is expected that a 35 subtilisin analog according to the present invention possesses superior properties to wild type subtilisins

- 44 -

in detergent formulations such as those disclosed in, for example, U.S. Patent No. 3,732,170; U.S. Patent No. 3,749,671 and U.S. Patent No. 3,790,482, all of which are incorporated by reference herein.

5 Moreover, for practical reasons many industrial processes are conducted at temperatures that are above the stability range of most enzymes. Therefore, although detergent applications have been emphasized herein, it is believed that thermostable 10 proteases according to the present invention are not only advantageous to certain industries such as detergent and hide dehairing, which already require stable proteases, but also may be useful in industries that use chemical means to hydrolyze proteins, e.g. 15 hydrolysis of vegetable and animal proteins for the production of soup concentrates.

Therefore, it is intended that the present invention be include all such modifications and improvements as come within the scope of the present 20 invention as claimed.

25

30

35

ANNEX M3

International Application No: PCT/US87/00027

## MICROORGANISMS

Official Sheet in connection with the microorganism referred to on page 13, line 15 of the description.A. IDENTIFICATION OF DEPOSIT: Escherichia coli containing plasmid pCFM936  
Further deposits are identified on an additional sheet 

Name of depositary institution: American Type Culture Collection

Address of depositary institution (including postal code and country):

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit:

13 January 1986

Accession Number:

ATCC 53413

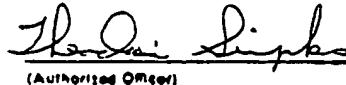
B. ADDITIONAL INDICATIONS: (Leave blank if not applicable). This information is continued on a separate attached sheet 

"In respect of those designations in which a European patent is sought, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 23(4) EPC)."

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE: (If the indications are not for all designated States)

D. SEPARATE FURNISHING OF INDICATIONS: (Leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later. (Specify the general nature of the indications e.g. "Accession Number of Deposit")

E.  This sheet was received with the international application when filed (to be checked by the receiving Office)


(Authorized Officer)

 The date of receipt (from the applicant) by the International Bureau is

(Authorized Officer)

- 46 -

WHAT IS CLAIMED IS:

1. An analog of a Bacillus subtilisin, said Bacillus subtilisin having an amino acid sequence comprising an Asn-Gly sequence, wherein one or both residues of said Asn-Gly sequence are deleted or are replaced by a residue of a different amino acid.
2. The analog as recited in claim 1 wherein an asparaginyl residue in said Asn-Gly sequence is replaced by a residue of a different amino acid.
3. The analog as recited in claim 1 wherein an asparaginyl residue in said Asn-Gly sequence is replaced by a residue of an amino acid from the group consisting of serine, valine, threonine, cysteine, glutamine and isoleucine.
4. The analog as recited in claim 3 wherein said asparaginyl residue in said Asn-Gly sequence is replaced by serine.
5. An analog of a Bacillus subtilisin having an amino acid sequence wherein an asparaginyl residue naturally occurring at a position corresponding to position 218 in the amino acid sequence disclosed in Table 1 is replaced by a residue of a different amino acid.
6. The analog as recited in claim 5 wherein an asparaginyl residue at said position is replaced by a residue of an amino acid from the group consisting of serine, valine, threonine, cysteine, glutamine and isoleucine.

- 47 -

7. The analog as recited in claim 6 wherein a seryl residue replaced said asparaginyl residue.

8. The analog as recited in claim 5 wherein 5 said analog is an analog of a naturally occurring Bacillus subtilisin in a strain selected from the group consisting of subtilisin Carlsberg, subtilisin DY, subtilisin BPN' and aprA gene product.

10 9. The analog as recited in claim 8 wherein all residues except the residue at position 218 comprise amino acid residues as disclosed in Table 1.

10. A nucleic acid encoding an analog of 15 Bacillus subtilisin, said Bacillus subtilisin having an amino acid sequence comprising an Asn-Gly sequence, wherein codons which encode one or both residues of said Asn-Gly sequence are deleted or are replaced by codons encoding a different amino acid residue.

20

11. The nucleic acid as recited in claim 10 encoding a polypeptide wherein an asparaginyl residue at a position corresponding to position 218 in the amino acid sequence disclosed in Table 1 is replaced by a 25 residue of a different amino acid.

12. The nucleic acid as recited in claim 11 wherein said different amino acid is a member of the group consisting of serine, valine, threonine, cysteine, 30 glutamine and isoleucine.

13. The nucleic acid as recited in claim 12 wherein said residue at said position comprises a seryl residue.

35

- 48 -

14. The nucleic acid as recited in claim 13 wherein all amino acid residues at all positions in said amino acid sequence of said analog except position 218 comprise amino acid residues as disclosed in Table 1.

5

15. A system for the production of an analog of a Bacillus subtilisin comprising:

a host cell; and

10 nucleic acid, within said host cell, encoding an analog of subtilisin having an amino acid sequence wherein a residue at a position corresponding to position 218 in the amino acid sequence disclosed in Table 1 does not comprise an asparaginyl residue.

15

16. The system as recited in claim 15 wherein said residue at said position comprises a residue of an amino acid from the group consisting of serine, valine, threonine, cysteine, glutamine and isoleucine.

20

17. The system as recited in claim 16 wherein said residue at said position comprises a seryl residue.

25

18. The system acid as recited in claim 17 wherein all residues at all positions except 218 comprise amino acid residues as disclosed in Table 1.

30

19. The system as recited in claim 18 wherein said host cell is deficient in secreted proteases other than subtilisin.

20.

The system as recited in claim 15 wherein said nucleic acid is extrachromosomal.

35

21. The system as recited in claim 18 wherein said nucleic acid is chromosomal.

- 49 -

22. A method for improving the thermal and pH stability of a Bacillus subtilisin having an amino acid sequence comprising an Asn-Gly sequence comprising the step of replacing a residue in said Asn-Gly sequence 5 with a residue of a different amino acid.

23. The method as recited in claim 22 wherein said residue is the asparaginyl residue at the position in the amino acid sequence of the Bacillus subtilisin 10 which corresponds to position 218 in the amino acid sequence as disclosed Table 1.

24. A method for improving the thermal and pH stability of a Bacillus subtilisin having an amino acid sequence comprising an Asn-Gly sequence comprising the step of deleting a residue in said Asn-Gly sequence. 15

25. A detergent formulation comprising an analog of a Bacillus subtilisin, said Bacillus subtilisin having an amino acid sequence comprising an Asn-Gly sequence, wherein one or both residues of said 20 Asn-Gly sequence are deleted or are replaced.

26. An analog of a Bacillus subtilisin, said 25 analog having an amino acid sequence with fewer Asn-Gly sequences than are present in said Bacillus subtilisin.

FIG. 1

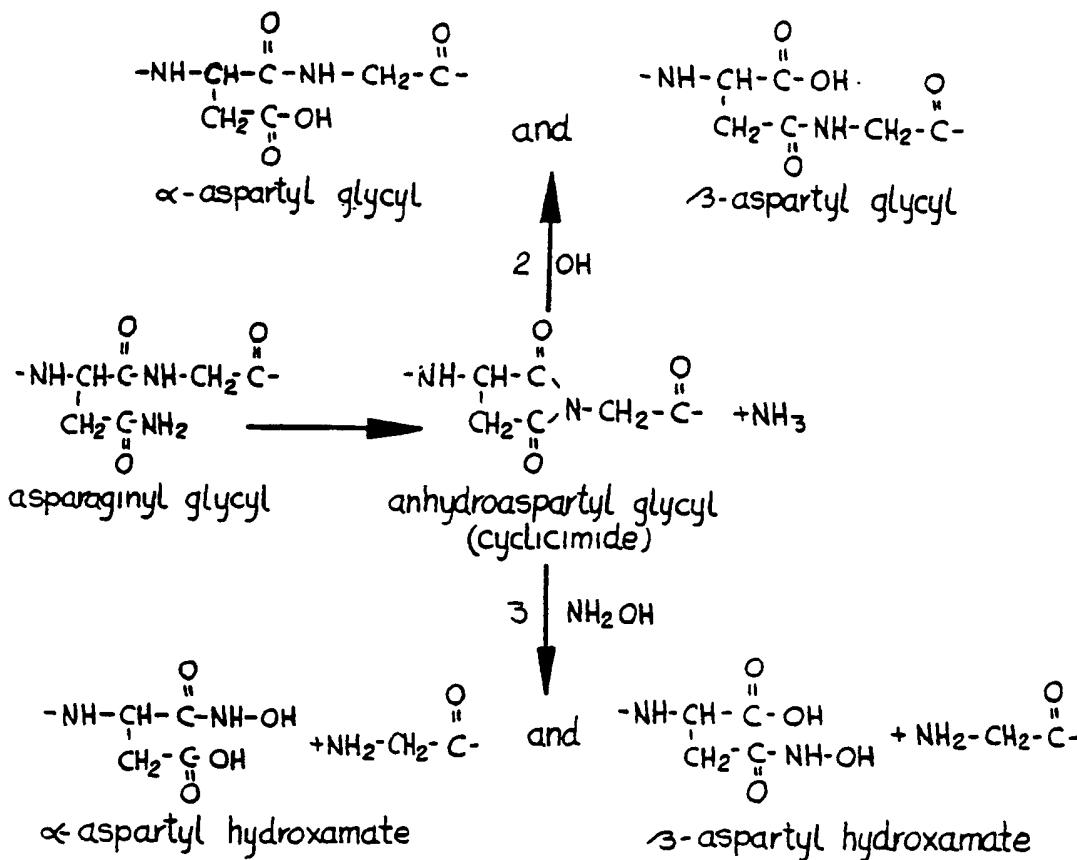
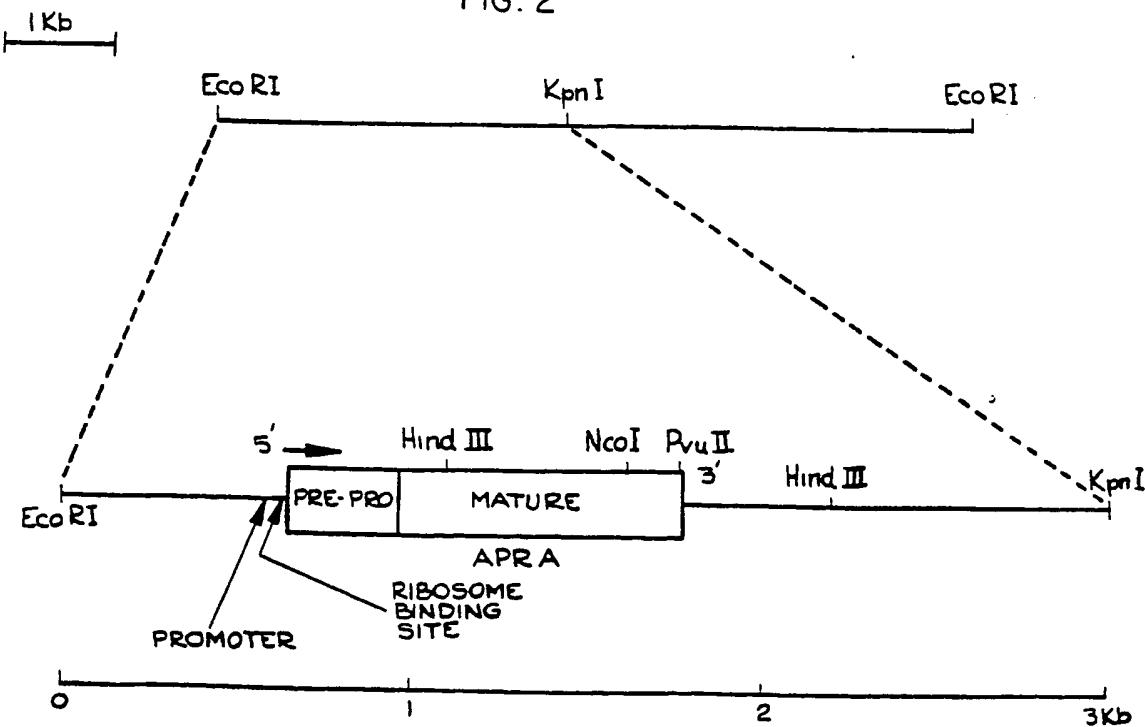


FIG. 2



2 / 3

FIG. 3

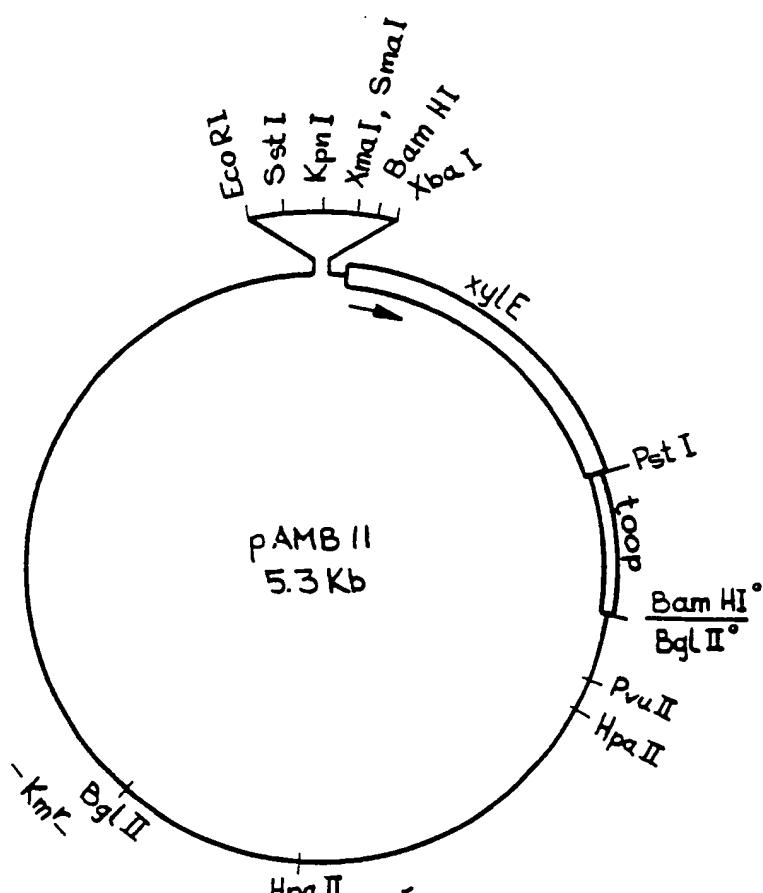
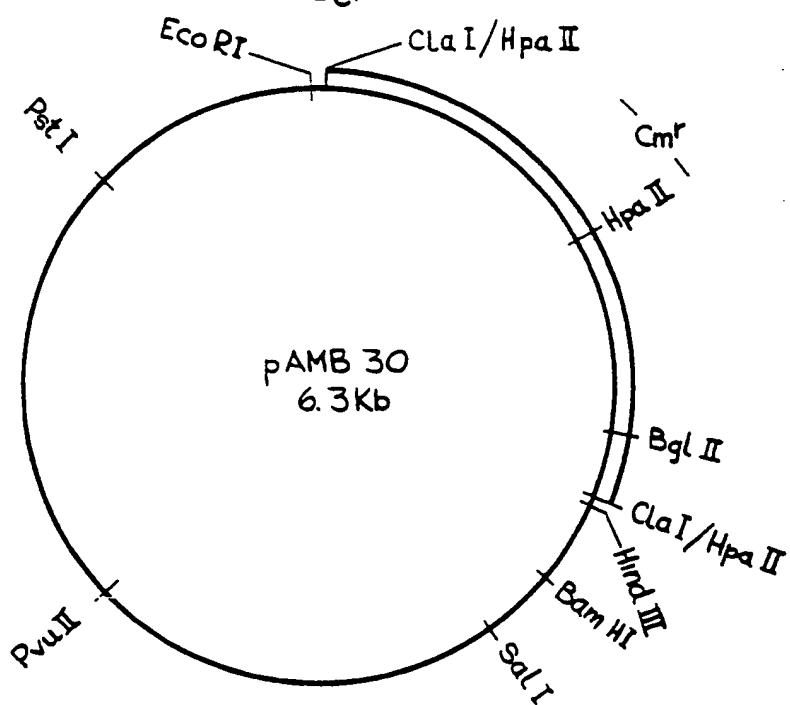


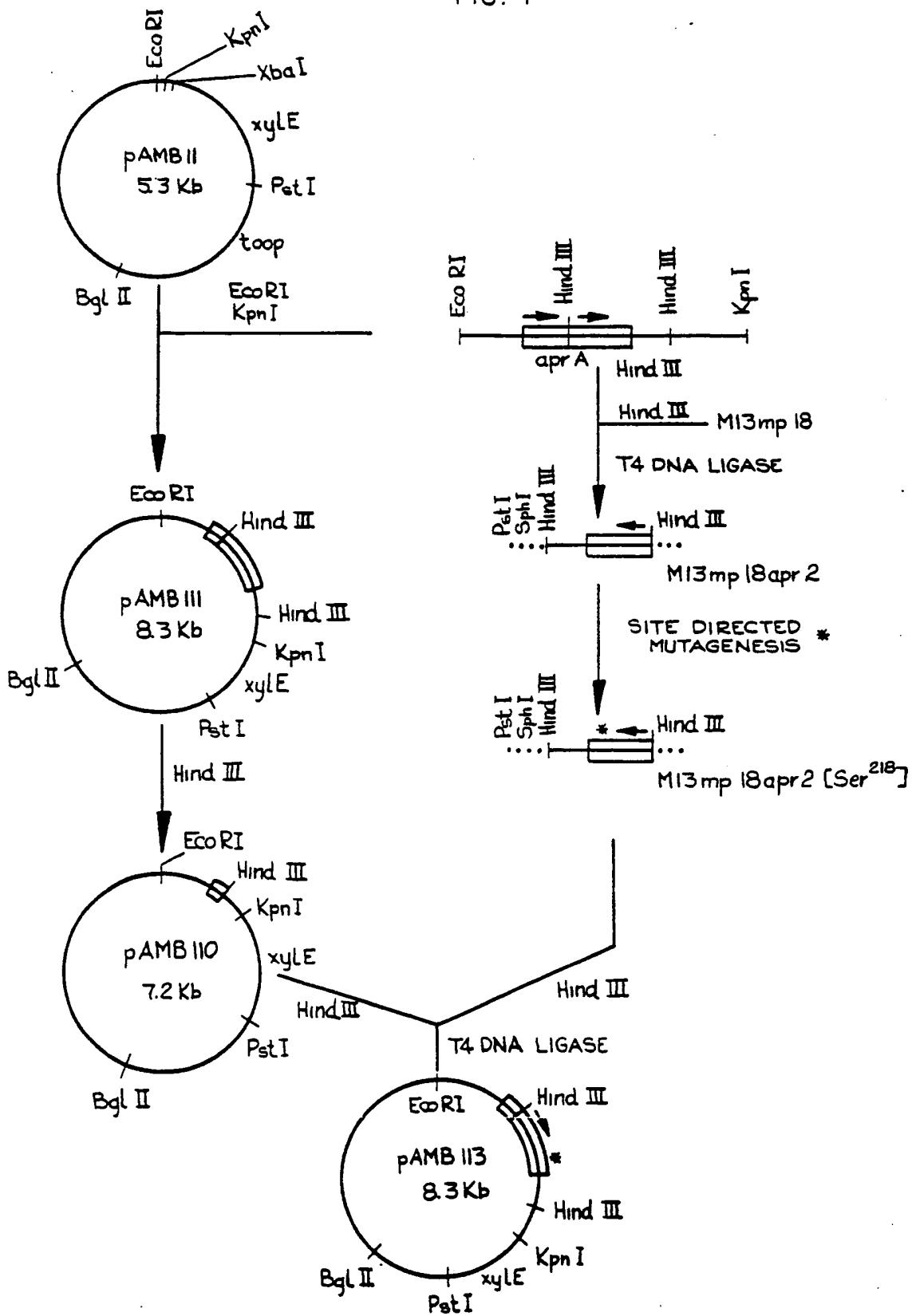
FIG. 5



□ pBD64 DNA

— pBR322 DNA

FIG. 4



# INTERNATIONAL SEARCH REPORT

International Application No PCT/US87/00027

## I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) <sup>3</sup>

According to International Patent Classification (IPC) or to both National Classification and IPC  
 IPC(4) C12N 9/56; C07H 15/12; C12P 21/02, 15/00 C12N 1/20  
 U.S. 435/222; 536/27; 435/70, 172.3 and 253

## II. FIELDS SEARCHED

Minimum Documentation Searched <sup>4</sup>

Classification System	Classification Symbols
U.S.	435/70, 172.3, 221, 222 and 253 536/27 935/14

Documentation Searched other than Minimum Documentation  
to the Extent that such Documents are Included in the Fields Searched <sup>5</sup>

**CHEMICAL ABSTRACTS DATA BASE (CAS) 1967-1987, BIOLOGICAL ABSTRACTS DATA BASE (BIOSIS) 1967-1987 See continuation.**

## III. DOCUMENTS CONSIDERED TO BE RELEVANT <sup>14</sup>

Category <sup>6</sup>	Citation of Document, <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
Y	<u>JOURNAL OF BACTERIOLOGY</u> (Washington, D.C.) Volume 158, issued May 1984 (STAHL ET AL) "Replacement of the <u>Bacillus subtilis</u> subtilisin structural gene with an in vitro derived deletion mutation". See pages 412 and 414.	1-26
Y	<u>JOURNAL OF BACTERIOLOGY</u> (Washington, D.C.) Volume 159, issued September 1984 (VASANTHA ET AL) "Genes for alkaline protease and neutral protease from <u>Bacillus amyloliquefaciens</u> contain a large open reading frame between the regions coding for signal sequences and mature protein". See pages 811-813.	1-26
Y	<u>NATURE</u> (London) Volume 318, issued 28 November 1985 (THOMAS ET AL) "Tailoring the pH dependence of enzyme catalysis using protein engineering". See pages 375-376.	1-26

\* Special categories of cited documents: <sup>13</sup>

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

## IV. CERTIFICATION

Date of the Actual Completion of the International Search <sup>2</sup>

17 March 1987

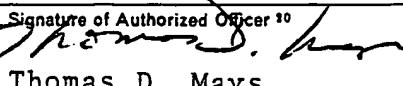
International Searching Authority <sup>1</sup>

ISA/US

Date of Mailing of this International Search Report <sup>3</sup>

27 MAR 1987

Signature of Authorized Officer <sup>10</sup>

  
Thomas D. Mays

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

Y	EP, A, 130,756 (BOTT ET AL) 01 September 1985. See pages 4, and 40-45.	1-26
Y	<u>METHODS IN ENZYMOLOGY</u> (New York) Volume XLVII, issued 1977 (BORENSTEIN ET AL) "Cleavage at Asn-Gly Bonds with hydroxylamine". See pages 134-137.	1-26
Y	US, A, 3,985,686 (BARRAT) 12 October 1976. See columns 7-9.	25 and 26
Y	US, A, 4,052,262 (HORIKOSHI ET AL) 04 October 1977. See columns 15 and 16.	25 and 26

VI.  OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE <sup>10</sup>

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1.  Claim numbers ..... , because they relate to subject matter <sup>11</sup> not required to be searched by this Authority, namely:

2.  Claim numbers ..... , because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out <sup>12</sup>, specifically:

VI.  OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING <sup>13</sup>

This International Searching Authority found multiple inventions in this international application as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4.  As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

- The additional search fees were accompanied by applicant's protest.
- No protest accompanied the payment of additional search fees.

PCT/US87/00027

II. FIELDS SEARCHED CONTINUED.

AUTOMATED PATENT SYSTEM (US) DATA BASE:

Keywords: *subtilisin, mutation, cloning,  
vector and plasmid.*

THIS PAGE BLANK (USPTO)